图书情报工作
第65卷 第23期 2021年12月

专家个体多维特征刻画与专家组均衡推荐研究*

■ 华斌　吴诺　贺欣
天津财经大学理工学院　天津 300222

摘　要：[目的/意义]提出一种基于专家个体多维特征刻画的电子政务项目评审专家组推荐方法，提升专家组间项目评审的一致性水平。[方法/过程]以专家个体的长期评审意见为数据源，利用意见挖掘技术实现知识元识别与情感极性获取；构造专家的领域知识结构并动态迭代更新；利用统计分析刻画专家知识水平、评审深刻性、情感风格、领域专长特征，实现基于科学计量的专家特征刻画并以此为基础进行专家组合的推荐。[结果/结论]本文的方法注重专家组的多维特征均衡，对电子政务项目评审具有很好的问题针对性，并在实践中取得了良好的应用效果。

关键词：电子政务　项目管理　意见挖掘　知识单元计量　专家推荐
分类号：G250
DOI：10.13266/j.issn.0252-3116.2021.23.007

1　引言

以认知论学说为基础的同行专家评价法[1]被我国政府部门广泛用来进行各类项目评审。同行专家评价法的优点是能够在缺乏足够统计数据和原始资料的情况下利用专家经验快速做出定量估计。专家评价结果的准确程度，主要取决于专家的阅历经验、知识丰富程度以及个人情感态度等因素。基于专家独立评价的专家组评价方法作为一种群决策方法，是对专家评价法的一种完善，但却又带来了专家组构成合理性的问题。在政府各类项目评审的实际工作中，随机抽取专家构成专家组的方式往往导致同一项目在不同组间的评审结论差异较大。因此，对专家组构成方法的研究不仅具有理论价值，而且具有提升项目评价科学性的实际意义。

科技项目具有多类性、技术路线多样性以及多目标性的个性化特点，如基础研究类与应用研究类的差异、物理实验与仿真实验的技术方法区别以及小试、中试、批量生产的不同阶段划分。理论上，科技项目评价的有效性主要取决于评价者的经验知识与评价目标的契合度，因此科技项目评价专家的推荐研究主要聚焦于专家知识结构与评价目标的一致性方面。区别于科技项目评价，电子政务项目评审具有更多的政策合规性、预算合理性、绩效考核可实现性等行政管理内涵。电子政务项目受到较为固定的管理规制约束，项目建设内容与建设结果又要遵守类属明确的技术标准与实施标准规范。因此，电子政务项目评审活动主要是以有限维度的确定性知识在有限时间内依赖专家经验进行的群决策行为，这就导致了电子政务项目评审专家组的知识结构合理性对于项目审批结论的科学性具有十分重要的支撑作用。基于认知科学理论和心理学理论，专家个体的多种特征差异是导致认知差异和情感极性差异的根本原因，也是造成专家组间评审结果不一致这一客观现象的来由。专家组评价作为一种旨在均衡专家个体差异的群体决策方法，其实质是应形成一套促进专家组多维特征均衡的推荐方法，以此提升专家组项目评审结论的一致性。据此，定义与刻画评审专家特征、实现基于特征均衡的专家组推荐研究对于电子政务项目审批工作具有重要的基础价值。

2　相关研究

专家推荐研究在科学学与情报学中均有涉及，但侧重不同。科学学的研究更加注重对专家研究专长与所评价项目的一致性、专家阅历与权威性等要素的选择，目的是尽可能实现同行评价、权威评价，解决复杂知识边界带来的认知差异问题。情报学的研究则侧重

* 本文系天津市信息化专项资金项目“电子政务项目全流程管理信息系统建设与软件项目预算评价系统”（项目编号：津党网信函（2018）146号）研究成果之一。
作者简介：华斌，教授，博士，E-mail：bigsoon@sina.com；吴诺，高级实验师，博士研究生；贺欣，硕士研究生。
收稿日期：2021-05-04　修回日期：2021-08-23　本文起止页码：58-69　本文责任编辑：易飞

华斌，吴诺，贺欣. 专家个体多维特征刻画与专家组均衡推荐研究[J]. 图书情报工作,2021,65(23):58-69.

对具有特定知识类型专家的发现，是科学学研究的重要工具。

清华大学[2]"人才评价测不准原理"研究成果中证明，对一个人各方面能力的评价需要通过科学地选择测试样本，并经过相当一段时间的过程性跟踪测试，才能获得对其学习和创造能力评价的信度和效度，且结果不是精确值。这一结论为本文利用长期的、历史形成的专家评审意见作为专家特征刻画的数据源提供了科学性支撑。同时，采用累积性、间隔性的动态刻画策略也完全符合专家认知成长的客观规律以及"过程性跟踪测试"的方法要求。中国人性格研究组将国人的性格分解为 5 个集质[3]，即：生活旨趣、认知风格、情绪特征、意志品质、态度倾向。其中，生活旨趣中包括知识性等特质；认知风格包括客观性、全面性、敏捷性等特质；情绪特征包括强烈性、持续性等特质。上述成果为本文的专家特征选择奠定了理论基础。

对于专家推荐研究，国外多从定性的角度探讨如何制定科学的同行评议专家监督管理制度[4-5]，国内则侧重于专家评价指标体系的构建以及专家识别推荐的方法：①在专家评价研究方面，游庆根等[6]提出一套由基本情况、科研能力、评审技能、个人信誉构成的专家评价指标体系，与陈媛等[7]以科研活跃程度、评审绩效和评审态度构成的评价指标体系具有不同的侧重考虑，为本文的专家特征选取提供了借鉴；②在专家推荐方法上，赵千等[8]结合论文子主题覆盖率与专家权威度两方面构建了专家推荐模型，王梓森等[9]按照项目学科树层次结构逐层计算专家学术专长与项目所涉及学科概念间的语义相似度，实现了多粒度的同行专家学术专长匹配方法，为本文的知识获取方法提供了启发。但是，上述研究大多缺乏对专家个体评价结论的知识计量分析，无法支撑利用元评价理论[10]进行专家组的推荐研究。

科学计量学与知识计量学是专家专长识别的基础理论。朱庆华[11]解析了知识元挖掘原理，辨析了知识元与知识单元的关系。进一步地，姜春林等[12]提出了利用知识单元计量实现学术论文评价的方案；贺颖、邱均平[13]利用知识图谱与科学计量方法实现交叉学科专家遴选。上述研究为专家知识特征定义提供了借鉴，也为本文的专家组构成理论提供了依据。针对专家评价问题，元评价理论提供了理解专家评价质量偏差的方法。元评价指标主要包括偏差系数、变异系数[14]等基于评分数据构建的客观指标，这是专家特征计量的基本方法论。

基于上述总结，在评价目标导向的专家特征定义基础上，以电子政务项目评审过程中累积的专家意见作为主要元评价内容，利用知识挖掘、知识计量的基本方法可以实现对专家个体的多维特征刻画，进而完成基于多维特征均衡的专家组推荐。

3 研究框架与相关方法

3.1 专家特征刻画的数据源与知识表示、补充策略

电子政务项目评审专家一般具有长期的评审经历，项目管理系统中不仅存有多个专家组的评审结论，而且保存了专家组中各专家的独立评审结果与专家意见短文本（见图 1）。上述数据作为专家个体特征刻画的客观依据，在情景知识辅助下即可满足专家组特征均衡推荐的应用需求。

xx市信息化项目专家论证意见	
	项目序号：xxxx
项目名称	xxxxx
申报单位	xx单位
一、项目的必要性和需求分析	
1. 项目建设的必要性	☐10分 ☐8分 ☑6分 ☐4分 ☐2分
2. 项目需求分析的明确性	☐10分 ☐8分 ☐6分 ☑4分 ☐2分
二、项目的主要内容和目标	
3. 项目目标的可行性	☐10分 ☐8分 ☐6分 ☑4分 ☐2分
4. 项目主要建设内容与目标的一致性	☐10分 ☐8分 ☐6分 ☑4分 ☐2分
5. 考核指标的合理性	☐10分 ☐8分 ☐6分 ☑4分 ☐2分
三、项目实施基础	
6. 技术路线或方案的可行性	☐10分 ☐8分 ☐6分 ☐4分 ☑2分
……	☐10分 ☐8分 ☐6分 ☐4分 ☐2分
总得分：42分	建议资金：0万元
总体评价意见	☒建议立项 ☑不宜立项
专家意见及建议： 1.项目建设方案不合理，软件和硬件运行平台方案设计没有分析，建设内容与已有信息化之间的关系没有说明；系统整体设计框架要详细说明内容。 2.UPS设备没有给出参考价格，另外，网络设备超出学校实际需求。 3.项目社会效益较好，但是考核指标缺少量化；项目需求分析需要细化。 专家签名：xxx　xx年xx月xx日	

图 1　电子政务项目专家评审意见表示例

注：来源于××市委网信办

专家意见作为专家对项目内容的认知结果，是知识单元的物化存储，其思想、观点蕴含于这些知识单元之中[11]。通过对专家意见短文本的挖掘，可以抽取出其中的知识元，从而形成知识计量的基础。知识单元的差异在于知识元与排列逻辑的不同，这是形成类比的基础。专家特征刻画的实质就是实现专家的类比。要完成这一类比，首先需要从类比源中抽象出某种隐含的属性，即完成知识抽象化；其次需要将“类比知识单元”结合某个待解决问题的实际进行知识创新，以形成类比后的新产物[15]。

具体到本文的问题，专家意见中的知识元挖掘要经过两个步骤：①完成知识单元的表达和组织，完成从物理层次的文本单元向认知层次的知识单元转换；②完成知识内容的计量，完成从语法层次向语义层次的转换。短文本意见挖掘的难点在于其文本特征稀疏和上下文缺失的特点，导致语义特征上存在逻辑包含关系等较为复杂的情况。要有效实现短文本意见挖掘，必须具有科学的知识补充与知识表示方法作为支撑。利用电子政务项目知识概念树与项目审批知识本体辅助完成对专家意见的知识获取与知识聚合，构成了本文实现专家特征刻画的技术方法。概念树的节点与本体中的实体等元素均由国家标准及对电子政务项目的相关管理政策进行标准化规范[16-20]。概念树作为目标知识语义深刻性的表示方法，可以解决专家意见中的概念包含关系语义识别与评审深刻性描述等问题。项目审批知识本体可以突破概念树的界限，将部分专家意见中涉及的、概念树无法映射的评审知识语义通过本体完成语义关联，进而发现隐性知识，扩展了知识计量的边界。知识本体也是项目背景知识的语义规范和知识关联依据。这种组合的知识表示方法为问题求解提供了很好的便利性。

3.2 研究框架

本文研究框架包含基于专家评审意见的知识获取、专家多维特征刻画、多维特征均衡的专家组推荐三个部分，如图 2 所示。其中，知识获取部分利用实体抽取技术进行实体语义识别，形成电子政务领域实体词库。在语法分析基础上细粒度拆分专家意见，提取 SAO（subject-action-object）[21]结构化知识。为达成知识分类、类比的目标，利用项目知识概念树有效识别专家意见中的知识层次并解决概念冲突问题，以及利用知识本体关联知识，挖掘隐性知识。该部分不仅在词汇层、句法层、语义层进行知识挖掘与获取；而且利用情感分类中被广泛应用的 RNN（循环神经网络）、LSTM（长短期记忆网络）及其对应的双向深度学习模型构建情感分类器进行训练，从而获取专家评审意见中的情感极性。

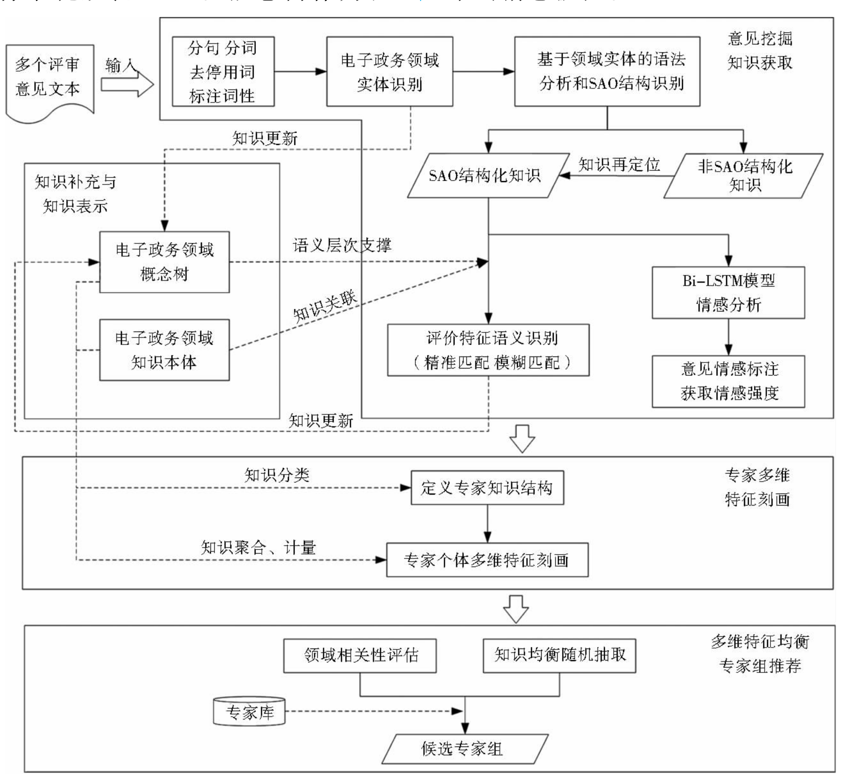


图 2 研究框架

华斌，吴诺，贺欣. 专家个体多维特征刻画与专家组均衡推荐研究[J]. 图书情报工作,2021,65(23):58－69.

首先使用开源工具 Hanlp 对评审意见文本初步分句、分词、词性标注、去除停用词;其次融入分词特征 X_i、词性特性 POS_i 提升模型识别效果,利用投票法集成 HMM、CRF、BILSTM、BILSTM-CRF4 种模型[22] 抽取出评审意见中所含实体;同时形成领域实体词库。

对多主语或多宾语并列的意见句进行拆分,利用依存句法分析获取 SAO 结构化知识,将其作为基础语义单元来表示专家意见。SAO 是一种从文本语料中抽取的三元组结构,在包含大量信息的同时,还可以有效地保持信息之间的内在关联性。在评审意见中,内在语义关联性主要包含 SA(主谓)、AO(谓宾)、SAO(主谓宾)三种,通常是专家对审批对象(多为主语或宾语形式)表明看法或调整建议。由于存在语言表述不规范的短文本意见,该部分语法分析效果不佳,导致 SAO 结构化知识抽取错误而无法识别文本中的内在语义关系。为解决这一问题,本文将评审意见中出现的主要问题归为三类,并针对性地自定义如下规则以再定位知识,其示例如表 1 所示:

表 1　专家审批意见非 SAO 结构化知识语义分析

原因	示例 注:知识抽取结果以“S/A/O”表示	自定义规则
①标定的依存关系过于复杂以及错误	1. 原句: “该项目方案包括将多个运维服务方案合成一个项目。” 知识抽取结果: 该项目方案/包括/合成项目	定义 1:原句进行分词、词性筛选,定位名词、动名词为概念候选词;定位动词、形容词为修饰性候选词,表示内在语义关系
②复杂语句存在两个核心依存关系	1. 原句: 合同复印件作为附件放到申报书。 知识提取结果: 合同复印件/作为/附件 /放到/申报书	定义 2:主语从句(或宾语从句)的 SAO 结构化知识作为复合句的主语(或宾语),补全复合句 SAO 结构化知识
③标定依存关系错误,名词性修饰语标定为并列宾语,而非谓语的定中关系	1. 原句: “减少信息化建设的重复投资。” 知识抽取结果: ① /减少/信息化建设 ② /减少/重复投资	定义 3:将两者的宾语合并为一个宾语,形成新 SAO 结构化知识,将其内在语义包含关系在知识中体现

3.3　专家多维特征选取

专家特征选取具有明确的目标约束性——电子政务项目审批知识约束。专家基于自身知识结构对项目做出评价,本质上是一种评价主体和客体间的知识交换活动[23]。电子政务项目作为政务职能建设的信息化工程,对专家知识结构的广度有更多要求。据此,以领域概念树与领域知识本体为参考,定义专家知识结构 4 个构成元素以进行知识分类,如表 2 所示:

表 2　专家知识结构定义

专家知识结构构成元素	所对应的项目要素
政务知识	需求、建设目标、政务职能、政策法规
技术知识	技术方案、建设内容
管理知识	实施方案、考核指标、效益分析
预算知识	经费预算

借鉴元评价理论与客观元评价指标[14] 的基本理论,定义知识水平、评审深刻性、情感风格、领域专长为专家多维特征刻画指标。这一定义综合考虑了专家知识能力和影响专家评审的心理因素以及相关信息的可获取性。专家特征的内在逻辑解释见表 3。

表 3　专家多维特征评价指标

特征维度	定义
知识水平	对于任一专家知识构成元素,计量个体专家评分意见、情感与专家组整体平均评分和意见情感的相对偏差[14],以表征专家在该知识构成元素下的知识水平
评审深刻性[24]	对于任一专家知识构成元素,计量个体专家意见中所涉及的领域特征词汇出现频次与概念树的分布层次、分布路径和集中度[24],以表征专家在该知识构成元素下的知识深度
情感风格	计量专家意见综合情感强度,以表征专家情感风格偏向
领域专长	专家的多次评审活动累计涉及多个内容主题,计量项目内容主题,以反映专家对相关领域的熟悉程度[8]

3.4　多维特征均衡的专家组推荐

上述专家特征选取针对电子政务项目管理目标,参考了已有的研究方法与数据的可获取性。以专家组间多维特征的计量结果为依据实现基于特征均衡的专家组推荐在理论上可以支撑本文目标的实现。在随机抽取的基础上,本文的推荐方法不仅考虑了候选专家组的领域专长与待审项目内容具有更好的贴近度;而且以专家库全部专家知识广度与深度计量结果的平均水平作为推荐依据。需要说明的是,在一定数量专家所构成的专家库基础上,本文的策略可以确保候选专家组的可产生性与非唯一性,支持管理者对专家组的筛选。

3.4.1　领域相关性评估

通过 LDA 模型获取待评审项目的主题概率,并与

专家领域计算主题相似度可以得到专家与待评审项目领域相关度 $Sim(V_{P_a},V_{E_i})$，该值越大，表明专家与待审项目的领域相关性越高，计算方法如公式（1）所示。其中 V_{P_a} 表示待评审项目 P_a 的主题特征向量，T_{t,E_i} 表示专家 Ei 的主题特征向量 V_{E_i} 的第 t 主题下的平均概率，计算方法如公式（2）所示，M 表示专家 Ei 所审项目数量。

$$Sim(V_{P_a},V_{E_i}) = \frac{\sum_{t=1}^{T}(T_{t,P_a} \times T_{t,E_i})}{\sqrt{(\sum_{t=1}^{T} T_{t,P_a}^2)(\sum_{t=1}^{T} T_{t,E_i}^2)}} \qquad 公式(1)$$

$$T_{t,E_i} = \frac{1}{M} \sum_{m=1}^{M} T_{m,t,E_i} \qquad 公式(2)$$

3.4.2　知识均衡随机抽取

专家组间知识水平的均衡实际上是通过对组内各专家的知识聚合来实现的。具体方法为：①首先计算专家库中各专家针对表 2 中 4 个维度知识元素的知识水平与评审深刻性均值，记为 ε_{l_i} 与 ε_{d_i}（$i \in$ {政务知识，技术知识，管理知识，预算知识}）作为阈值；②将上一阶段得到的领域相关度作为权重 ω 融入随机抽取方法中；③从专家库（M 位专家）中随机抽取 m 位（$3 \leqslant m \leqslant 9 < M$，m 为奇数）专家构成候选专家组，分别计算候选专家组各知识结构的知识水平与评审深刻性均值，如果均超过阈值 ω_{l_i} 与 ω_{d_i}，保留本次抽取结果，否则重新进行随机抽取。

4　实证分析

4.1　数据源与实验环境

本文以 2017 年 – 2018 年 214 个省级电子政务项目专家组评审意见作为基础语料库，共计 1 211 条意见。领域本体基于 Cypher 语言存储于 Neo4j 中；实验均基于 Python 语言在 Windows 10 环境下实现。实验中利用 Google 开源深度学习框架 Tensorflow 及高级 API——keras 完成领域实体抽取；利用开源工具 Hanlp 提供的 pyhanlp 包对语句进行依存句法分析并完成结构化知识获取；利用 Echarts 工具软件进行结果的可视化呈现。

4.2　构建项目知识概念树与项目审批知识本体

4.2.1　电子政务项目知识概念树

根据电子政务项目评审指标要求以及知识结构定义，对本文的项目知识概念树做出以下约束：

（1）概念树的根节点为项目建设方案总概念，称为第 0 层概念；

（2）概念树的第 1 层概念为电子政务项目建设的八大要素：需求、建设目标、技术方案、建设内容、实施方案、考核指标、效益分析、经费预算；

（3）概念树中节点的深刻性权重 w_{ij} 的约束条件：子特征概念的深刻性权重大于父特征概念深刻性权重[24]。

在专家指导下，自顶向下划分概念层次、提取概念，依据哈尔滨工业大学同义词词林扩展版，合并同义词并人工调整。领域概念树如图 3 所示，$c_0 - c_{32}$ 表示概念，c_0 为树的根节点，syn 表示与之对应概念的同义词，从子概念指向父概念的箭头表示父子概念的层次关系[24]。概念树的广度与根节点概念的内涵丰富性正相关，而深度则与语义内涵的针对性正相关。遍历概念树则可实现对专家意见中不同概念层次的语义识别。

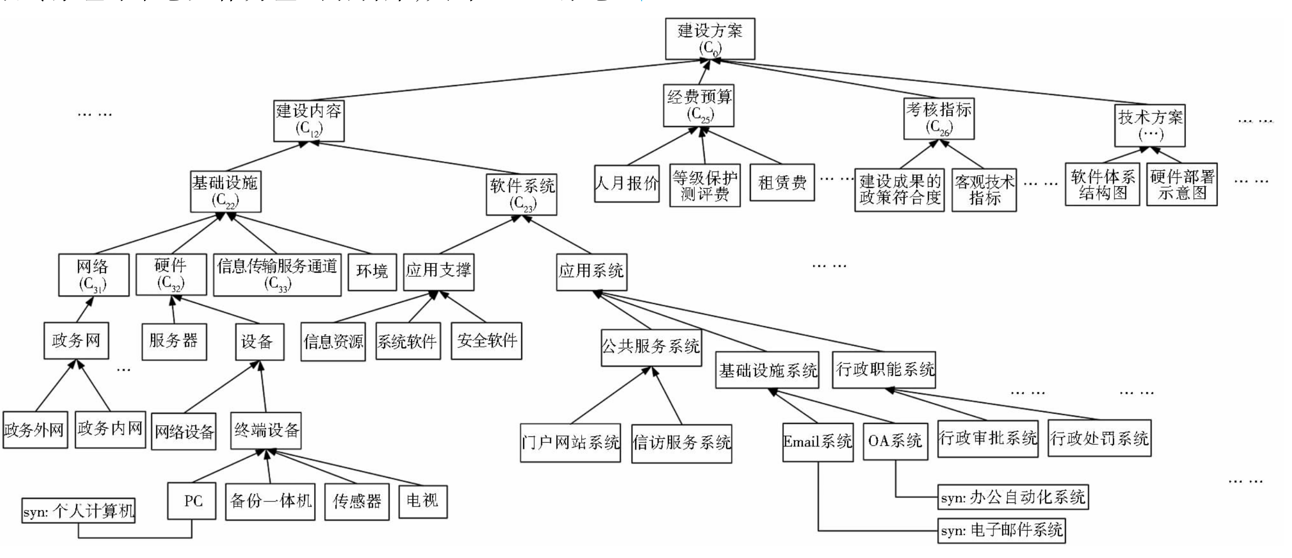


图 3　电子政务项目领域知识概念树示例（部分）

华斌，吴诺，贺欣. 专家个体多维特征刻画与专家组均衡推荐研究[J]. 图书情报工作,2021,65(23):58－69.

4.2.2 电子政务项目审批知识本体

电子政务项目审批知识本体(见图 4)对于背景知识的表示则更加完整,通过关联专家与项目之间的知识元,获取专家评审依据——政务职能和政策法规隐性知识,补充专家知识结构对应的项目要素内容,完善知识分类与聚合。该本体包括电子政务项目、评审意见、审批环节、政策法规、政务职能五大类,其中电子政务项目包括项目要素子类,政务职能包含政务子职能子类。表 4 定义了本体中概念间的语义关系;以政务职能知识为例,公式(3)定义了专家意见通过项目要素关联政务职能的知识关联规则。后续将关联得到的政务职能知识与政策法规知识分类到政务知识中(见表 2),补充知识结构中的项目背景知识。

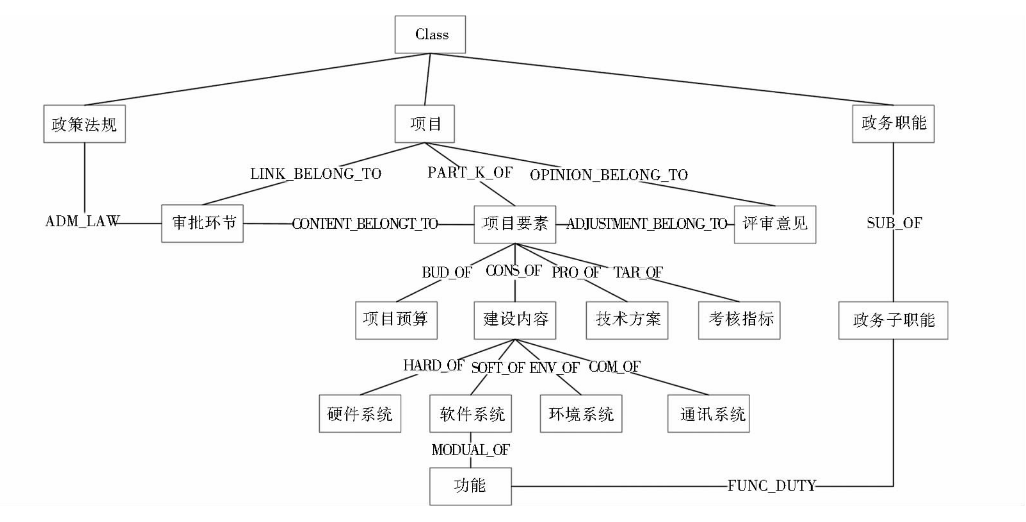


图 4 电子政务项目审批知识本体

表 4 概念间语义关系描述

通用关系名	语义关系描述
LINK_BELONG_TO(属于审批环节)	项目与审批环节之间的关系
ELEMENT_OF(包含项目要素)	项目与项目要素之间的父子类关系
OPINION_BELONG_TO(属于项目)	评审意见与项目之间的关系
ADM_LAW(属于审批环节)	政策法规与审批环节之间的关系
CONTENT_BELONG_TO(属于审批环节)	项目要素与审批环节之间的关系
ADJUSTMENT_BELONG_TO (属于项目要素)	评审意见与项目要素之间的关系
FUNC_DUTY(属于子职能)	项目要素与政务子职能之间的关系
SUB_OF(包含政务子职能)	政务职能与子职能之间的父子类关系
BUD_OF(包含项目预算)	项目要素与项目预算之间的父子类关系
CONS_OF(包含建设内容)	项目要素与建设内容之间的父子类关系
PRO_OF(包含技术方案)	项目要素与技术方案之间的父子类关系
TAR_OF(包含考核指标)	项目要素与考核指标之间的父子类关系
HARD_OF(包含硬件系统)	建设内容与硬件系统之间的父子类关系
SOFT_OF(包含软件系统)	建设内容与软件系统之间的父子类关系
ENV_OF(包含环境系统)	建设内容与环境系统之间的父子类关系
COM_OF(包含通信系统)	建设内容与通信系统之间的父子类关系
MODUAL_OF(包含功能)	软件系统与功能之间的父子类关系

$$\begin{cases} Opinion(a) \\ ADJUSTMENT_BELONG_TO(Element,\ Opinion) \\ FUNC_DUTY(Element,\ Subfunction) \end{cases}$$

$$\Rightarrow Opinion_Func(\exists ADJUSTMEBT_BELONG_TO(Opinion(a), Element).Element, Subfunction) \cap \forall Opinion_Func.Subfunction \qquad 公式(3)$$

其中 a 为某环节的专家评审意见 SAO 结构化知识变量,通过 *ADJUSTMENT_BELONG_TO* 关系获取评审意见 a 所关联的项目要素,通过 *FUNC_DUTY* 关系获取其项目要素所属的子职能实体。

最终,评审意见 a 与政务职能之间的知识关联结果表示为 $Opinion_Func(\exists ADJUSTMENT_BELONG_TO(Opinion(a), Element).Element, Subfunction)$, $\forall Opinion_Func.Subfunction$,表示该关系下的所有政务子职能。

4.3 专家评审意见挖掘与知识获取

利用集成模型[22]抽取专家评审意见文本中的领域实体,通过人工筛选去重,共获得 246 个准确领域实体,同义词扩展后形成领域实体库,共计 324 个。

在领域实体词库的基础上,利用 Hanlp 分析专家意见句法与内在语义关系,细粒度拆分专家意见,提取 SAO 结构化知识,完成知识分类。其中主语 S 和宾语 O 主要表示被评审对象,一般为名词、动名词等,这些词或词组与项目概念语义紧密相关。因此,针对主语和宾语,采用精准匹配和基于文本最大相似度的模糊匹配方法实现基于概念树的语义识别,在得到针对性强的细粒度概念层次的同时,能够归结到概念树第一

层概念节点下以完成知识分类。

综合专家个体意见与专家组综合意见，共获取1 752个SAO结构化知识。在专家指导下人工预先标注评审特征语义概念，利用标注结果与不同阈值的算法结果进行对比，在保证概念准确的条件下选择最大阈值，最终确定模糊匹配阈值为0.55。为有效检验语义识别结构，定义两种检验方法：①SAO结构化知识语义识别结果以准确率（P）、召回率（R）和F1值（F1）为指标；②专家组综合意见语义识别结果以有效率（U）为指标，计算单项目SAO结构化知识评审特征语义识别正确率（Project_S），其大于阈值的项目数量占比即为项目专家组综合意见语义识别结果有效率。根据项目专家评审意见数据集特点，在专家的指导下设定Project_S阈值为0.75，计算方法如公式（4）（5）所示。为有效检验知识分类结果，根据概念分布路径，分别将人工预先标注的评审特征语义概念与算法结果归结到概念树第一层概念节点下，以分类准确率（C_P）、分类召回率（C_R）和分类F1值（C_F1）作为实验指标。

$$Project_S = \frac{单项目中评审特征语义识别正确的SAO结构化知识数量}{单项目SAO结构化知识总数} \times 100\% \qquad 公式(4)$$

$$U = \frac{Project_S大于阈值的项目数量}{项目总数量} \times 100\% \qquad 公式(5)$$

表5 评审特征分词语义识别实验结果

SAO结构化知识语义识别结果			项目综合意见语义识别结果	SAO结构化知识分类结果		
P/%	R/%	F1/%	U/%	C_P/%	C_R/%	C_F1/%
94.18	86.01	89.91	81.42	97.08	88.49	92.58

结果如表5所示，其中有效率为81.42%，说明大部分电子政务项目专家组评审意见内容均可按照该方法挖掘评价特征语义。另外，SAO结构化知识语义识别结果的准确率P达到94.18%，F1达到89.91%，也有效证明了该方法可以较好地给评审特征赋予概念知识。对于知识分类结果，C_P、C_R、C_F1均略大于语义识别结果，由此可见，该方法虽然会因为概念语义的细粒度导致识别有误，但却不影响粗粒度的父概念知识分类。召回率R和C_R相对较低的原因是：①意见中存在无法识别的项目概念，需要利用领域本体才能识别；②少部分意见语句过于复杂，形成的SAO结构化知识语义不完整，可以利用表1自定义规则进一步识别；③概念树构建不完善，不可能前期完全覆盖项目知识的全部概念，可以对概念树进行动态扩充和维护，以提高召回率。

以评审意见SAO结构化知识为数据，以分类精度为实验指标。比较实验结果，RNN模型85.37%，LSTM模型86.79%，BIRNN模型89.15%，而BILSTM模型效果最好，精度达到90.09%。由此可见，选取BILSTM模型可以有效判定专家意见情感倾向。评审特征语义与情感强度部分结果如表6所示：

表6 知识结构

SAO结构化知识	评审特征	知识概念层级	顶层概念（项目要素）	情感倾向	情感强度
建设目标/明确/	建设目标	1	建设目标	正向	1.91
技术路线/不明确/	技术路线	2	技术方案	负向	-1.28
备份一体机数量/过多/	备份一体机	6	建设内容	负向	-1.47
项目设备调试费用/过高/	设备调试费用	2	经费预算	负向	-2.75
/没有/完整拓扑结构图	拓扑结构图	3	技术方案	负向	-1.83
业务需求/清晰合理/	业务需求	2	需求	正向	1.96
考核指标/需要/具体化	验收指标	1	考核指标	正向	1.02
部分功能建设/具有/必要性	功能建设	4	建设内容	正向	3.6
水质统计分析/细化/方案	水质统计分析	2（本体映射为子职能-水资源管理，自定义层级为2）	政务职能	负向	-0.81

4.4 个体专家多维特征刻画与可视化展示

4.4.1 知识水平与评审深刻性

（1）知识水平。依据表3中知识水平定义，该测度指标设计具体如下：

$$Level_t = 1 - \frac{1}{n}\sum_{j=1}^{n}(\alpha_1 * D_j + \beta_j * O_j), t \in \{政务知识,技术知识,管理知识,预算知识\} \qquad 公式(6)$$

$$D_j = \frac{|x_j - \bar{X}_j|}{\bar{X}_j} \qquad 公式(7)$$

$$O_j = \left| \frac{1}{K_1}\sum_{k_1=1}^{K_1} s_{k_1 j} - \frac{1}{K_2}\sum_{k_2=1}^{K_2} s_{k_2 j} \right| \Big/ \frac{1}{K_2}\sum_{k_2=1}^{K_2} s_{k_2 j} \qquad 公式(8)$$

华斌，吴诺，贺欣. 专家个体多维特征刻画与专家组均衡推荐研究[J]. 图书情报工作,2021,65(23):58－69.

其中 $\bar{X}_j = \frac{1}{m}\sum_{i=1}^{m} x_{ij}$，m 表示本次评审活动的专家组人数。

$Level_t$ 为在知识结构 t 元素方面专家的知识水平指标，通过权重 α_l 和 β_l 控制打分偏差和意见偏差的重要性。n 表示与该知识构成元素所对应的项目要素个数。为打分偏差系数[14]，表示对于项目要素 j，专家打分 xj 与所有专家均分 $\bar{X}_j$ 的相对偏差。O_j 为意见偏差系数，表示对于项目要素 j，个体专家的意见平均情感强度与专家组整体平均情感强度的相对偏差，s_{k1j}、s_{k2j} 分别表示个体专家、专家组对于项目要素 j 的意见情感强度，K_1、K_2 分别是属于项目要素 j 的个体专家与专家组意见数量。

（2）评审深刻性。专家越具有深刻性的评审意见，越能体现出其掌握电子政务项目建设所需的实际知识。因此，基于项目知识概念树语义层次性的专家评审深刻性指标及其测度方法，可以刻画专家的知识深度。

专家评审深刻性度量从以下方面进行描述[24]：①如果专家意见中所涉及的电子政务领域特征词汇量大，则意见内容可能较丰富和深刻，专家知识广度与深度可能较大；② 专家意见中某些特征词出现的频次较高，则专家对该特征所涉及的概念较为明确，理解较深；③特征词分布于概念树的层次、分布路径和集中度（节点的出度）决定了所评审特征语义内容的具体、明确程度，所涉及的知识结构元素具有针对性、突出性，或知识构成有所侧重。

其中 PF 为评审特征集合，RF_t 和 RF_t-count 分别为意见中属于某一知识结构的评审特征集合及其评审特征数，rf-count$_i$ 为评审意见中第 i 个评审特征的数量。

定义 1：专家评审意见中属于知识结构 t 元素的评审特征 k 的频数为：

$$rc_t - fre_k = \frac{rf_t - count_k}{\sum_{t=1}^{RF_t - count} rf_t - count_t}, \ t \in \{政务知识, 技术知识, 管理知识, 预算知识\} \qquad 公式(9)$$

定义 2：评审特征 c 的集中度为评审特征 c 的下位词 c-son 在评审意见中出现的次数与下位词集合元素的数量之比。因此，给出专家意见中评审特征 k 的集中度如下：

$$rc_t - center_k = \frac{|rf_t - son_1|}{|rf_t - son_2|} \qquad rf_t - son_1 \in RF_t、rf_t - son_2 \in PF \qquad 公式(10)$$

定义 3：评审深刻性为意见中所有评审特征词汇在领域特征概念树中出现的次数、集中度所反映出来的深刻性之和。其计算公式如下：

$$RDep_t = \alpha \times \sum_{k=1}^{RF_t - count} rc_t - fre_k \times w_k + \beta \times \sum_{k=1}^{RF_t - count} rc_t - center_k \times w_k \quad rc_{tk \in RF_t} \quad \alpha_d + \beta_d = 1 \qquad 公式(11)$$

（3）可视化展示与分析。设定两种偏差权重 α_l、β_l 为 0.5，综合平均专家每次评审的知识水平，得到专家整体知识结构水平；通过概念树语义层次识别对专家所有评审意见进行深刻性计算得到专家评审深刻值；经领域专家确定，针对上下位父子关系，将子概念节点的权重确定为其对应父概念节点权重的 1.2 倍，并设定两种深刻性度量权重 α_d、β_d 均为 0.5。为清晰了解同一名专家知识广度与深度，在同一张雷达图中呈现二者的知识结构分布情况。以专家 A、专家 B 为例，进行说明与分析。专家 A 知识结构的水平和评审深刻性如图 5 所示，从图中可以看出，专家 A 的知识水平与评审深刻性一致性较强。由图 6 可见，专家 A 的各项知识水平和评审深刻性均高于专家库均值，评审能力较为突出。而由图 7 可知，专家 B 的技术知识和管理知识水平低于专家库均值，相应地其评审深刻性也较低。可见，专家 B 的知识结构不均衡，能力偏向性较强。

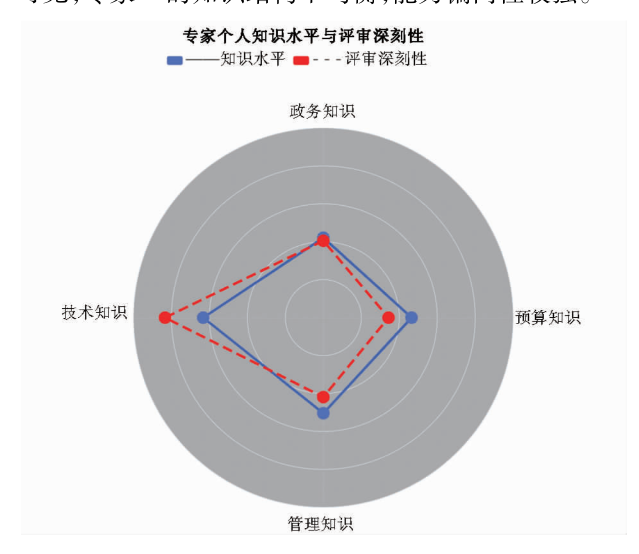


图 5　专家 A 的知识水平与评审深刻性

4.4.2　情感风格

专家情感通过评审意见传递，依据表 3 中情感风格定义，专家意见总情感强度 REmo 的计算方法如公式(12)所示，其中 K 是专家已审项目的所有意见总数，sk 表示专家意见 k 的情感强度。

$$RE_{mo} = \frac{1}{K}\sum_{k=1}^{K} s_k \qquad 公式(12)$$

以柱状图的可视化方式呈现多名专家评审情感风格，如图 8 所示。可见专家 X1、X4、X9、X11、X16、X19

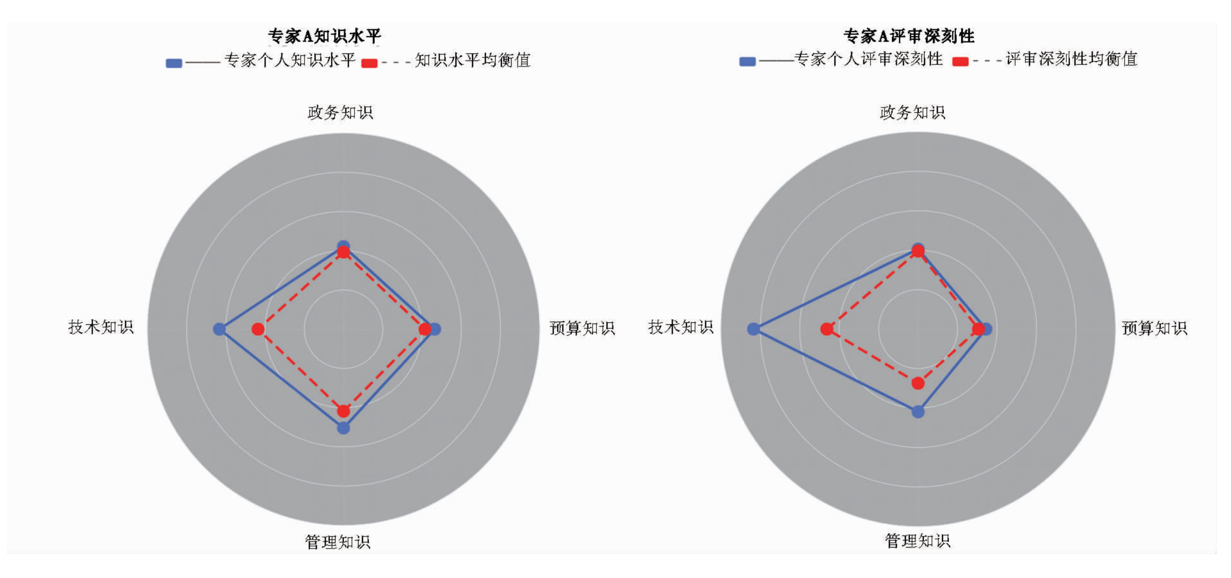


图 6 专家 A 知识水平、评审深刻性与专家库均值对比

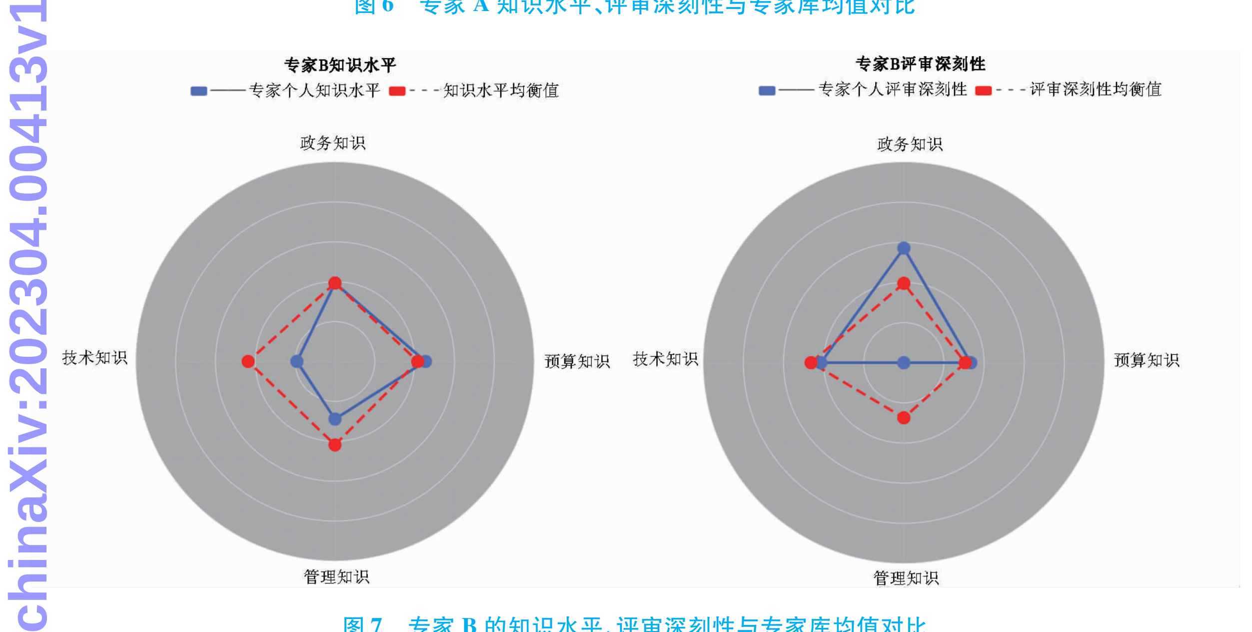


图 7 专家 B 的知识水平、评审深刻性与专家库均值对比

负情感较为强烈,言语较为直白,以多直接批评调整的评审风格为主;专家 X2、X5、X7、X8、X18 正情感与负情感较一致,言语较委婉积极,以多鼓励肯定项目建设的评审风格为主、以修改调整性意见为辅,对项目建设进行指导。核实专家原始意见,也符合上述情感风格的描述。

4.4.3 领域专长

由于 LDA 主题模型方法成熟且有效,本文选用该方法,以所有评审项目题目、材料摘要作为数据源,对被评审项目进行内容主题分析,以此反映专家所涉及的项目领域。主要步骤如下:① 进行数据清洗、去除标点符号、数字,过滤停用词;去除常用的软件描述用词,以提升 LDA 对于政务、项目功能、职能的主题表征度(如:模块、业务、平台、系统等);完成词袋构建。② 利用困惑度对数据集确定主题个数提升模型效果,同时利用 LDA 模型获得存有主题与词以及文档与主题之间关系的矩阵。③ 通过对该矩阵数据进行解析,获得每个主题下词的分布情况以及每个文档所属主题的情况,最后统计每位专家各已审项目所属主题的概率,分析专家的过往评审内容主题。

计算困惑度并通过实验结果多次迭代调整确定主题个数为 70,经 LDA 模型内容主题分析后,统计专家每一个已审项目的最大概率主题,进一步对主题词与权重进行词云展示。以专家 A 为例,已审 26 个项目,其中 4 个项目属于“学生教师心理救助辅导”主题,4 个项目属于“案件预警视频指挥”主题,4 个项目主题属于“案件执法监督与文书查询”主题。对频数较大的主题词进行词云展示,如图 9 所示,可以明显看出专家 A 评审领域偏向于“案件”“预警”“视频”等电子政务项目的建设。

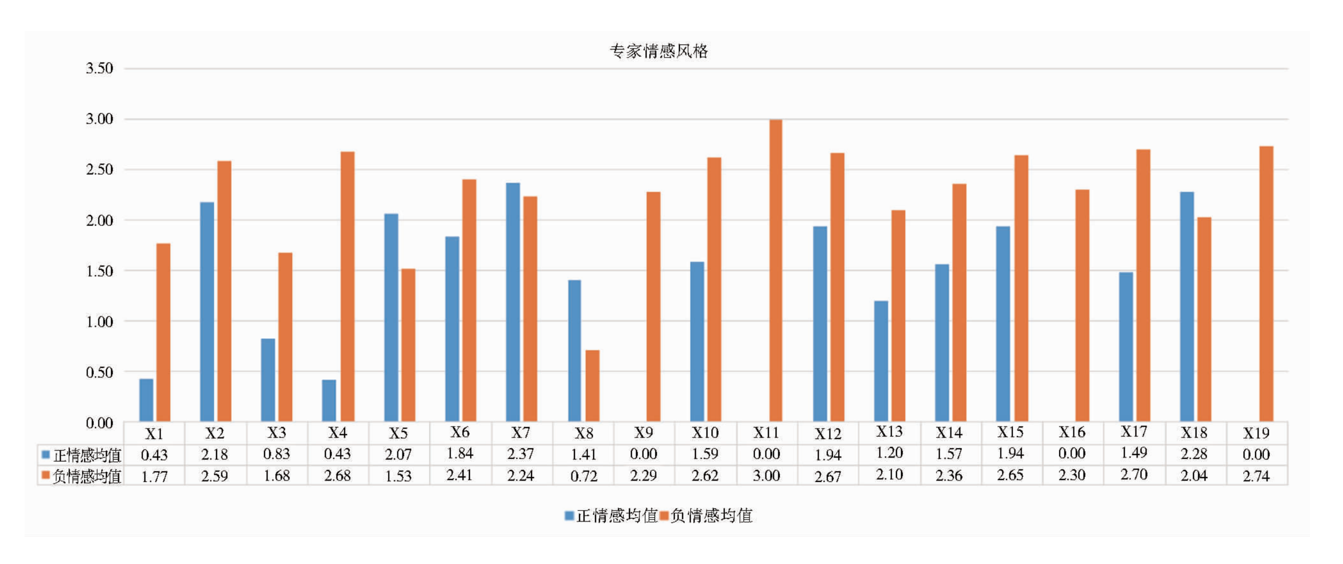


图 8　专家情感风格



图 9　专家 A 评审领域词云图

4.5　专家组多维特征均衡推荐结果

在实例化前述专家特征刻画方法的基础上，实现基于多维特征均衡的专家组推荐。在实际应用中，需要充分考量多项目同时段评审、专家人才库的有效使用率、专家精力等因素。为验证本文抽取方法在实际应用中的有效性，在单一项目专家组推荐验证的基础上，增加多项目并行的专家组推荐验证。选取网信专家库中的 50 位专家，设定专家组抽取人数为 5 人。

4.5.1　单一项目专家组多维特征均衡推荐

为验证多维特征均衡的单一项目专家组推荐结果，利用程序进行 100 次仿真抽取，结果如表 7 所示，后两列为被抽取的专家编号。在相同情况下可视化比较纯随机抽取与多维特征均衡抽取结果，直观展示方法的有效性。

以第 006 次抽取为例，经过多维特征均衡抽取/纯随机抽取得到的候选专家组知识水平均值与评审深刻性均值如图 10 短虚线/点虚线所示，专家库知识水平与评审深刻性均值如图 10 实线所示。可以看到，均衡推荐的专家组整体在 4 个知识构成元素上的知识水平均值与评审深刻性均值都等于或高于专家库均值，达到了平均水平，而采用纯随机推荐的专家组在技术、管理、预算 3 个知识构成元素上的知识水平均值均低于专家库均值，虽然该组评审深刻性在管理知识上高于专家库均值，但其他三个知识构成元素上都较大程度低于专家库平均标准。更进一步，分别列出两组候选专家的 5 位专家详细数据，如表 8 所示。结合表 8 可以分析得出，经过均衡抽取得到的 6 号、15 号专家在技术知识上具有较高水平和深刻性，弥补了 16 号专家在该知识上的不足；6 号、13 号、15 号专家在管理知识上弥补了 5 号和 16 号专家的不足。专家之间实现了知识互补。相比之下，纯随机抽取到的 3 号、5 号、8 号、10 号、17 号专家各知识元素的水平与深刻性参差不齐，知识互补程度较低，组内的均衡度较差。由此可得，均衡抽取方法较纯随机抽取增加了对知识均衡与互补性的考虑。

表 7　抽取结果对比

抽取次数	纯随机抽取结果	多维特征均衡抽取结果
001	19, 8, 20, 11, 21	16, 42, 28, 35, 6
002	20, 21, 1, 6, 17	1, 2, 6, 4, 7
003	38, 48, 50, 39, 29	6, 19, 7, 13, 17
004	17, 11, 15, 13, 1	40, 25, 26, 24, 6
005	6, 34, 47, 49, 8	4, 15, 25, 44, 50
006	3, 8, 10,17,5	16, 5, 6, 15, 13
007	1, 2, 13, 22, 17	15, 37, 39, 35, 6
……	……	……

此外，应用纯随机函数需要在很长时期内才能使每位专家抽中的次数趋于均衡，而在较短时间内是无法实现的[25]。观察表 8 可以发现，融入评审领域相关度的均衡抽取不仅能够提升专家组成员与待评项目所涉及领

域的贴近度,而且并不关联专家已审项目数量,在一定程度上可以消除纯随机抽取带来的次数不均衡问题。

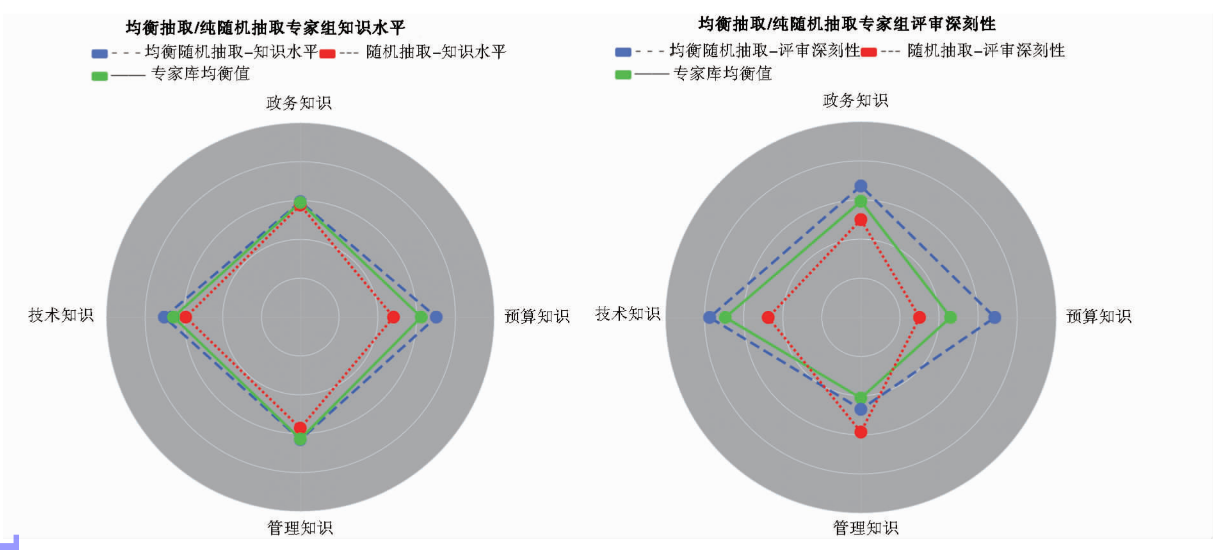


图 10 第 006 次均衡抽取/纯随机抽取专家组知识水平、评审深刻性对比

表 8 第 006 次均衡抽取/纯随机抽取专家组组内专家详情

抽取次数	专家编号	知识水平				评审深刻性				与待审项目的领域相关度	已审项目数
		政务	技术	管理	预算	政务	技术	管理	预算		
第 006 次均衡抽取专家组	5	0.650	0.827	0.499	0.715	0.500	0.543	0.500	0.500	0.749	7
	6	0.631	0.949	0.756	0.697	0.608	1.251	0.628	0.515	0.483	36
	13	0.628	0.539	0.784	0.584	0.531	0.614	0.500	0.500	0.484	8
	15	0.472	0.905	0.684	0.830	0.864	0.844	0.720	1.412	0.650	11
	16	0.590	0.287	0.433	0.681	0.857	0.621	0.000	0.500	0.496	3
第 006 次纯随机抽取专家组	3	0.511	0.734	0.690	0.260	0.500	0.706	0.720	0.500	0.794	3
	5	0.650	0.827	0.499	0.715	0.500	0.543	0.500	0.500	0.749	7
	8	0.586	0.389	0.496	0.461	0.500	0.000	0.500	0.000	0.087	3
	10	0.454	0.622	0.541	0.386	0.500	0.500	0.600	0.000	0.980	2
	17	0.683	0.384	0.633	0.581	0.500	0.625	0.610	0.500	0.389	4

4.5.2 多项目并行专家组多维特征均衡推荐

设定多项目并行抽取的约束条件:① 符合单一项目候选专家组多维特征均衡推荐;② 同时间段内,在有 M 名专家的专家库中抽取不同的候选专家组(组内 m 位专家),分别对 N 个项目进行评审;③ 不同待审项目中,不得出现同一名评审专家。

在上述条件约束下,用本文方法进行仿真实验,分别计算各专家与 N 个待审项目的领域相关度,在多维特征均衡的基础上进行抽取,得到多项目并行评审的最大项目数 Nmax = 6。该结果证明,本文方法在保证领域相关度和多维特征均衡的同时,能够充分满足同时段内多项目并行评审的实际应用。

5 结论与展望

利用知识计量、情感特征提取等技术手段进行专家个体多维特征刻画,并基于刻画结果进行专家组多维特征均衡的推荐方法从原理上进一步贴近了认知科学的理论内涵,较之以随机抽取组成的专家组具有更好的问题针对性与科学性。实践中,在较为明确的有限维度知识边界条件下,本文由于采用规范的知识补充与知识表示方法作为技术实现的基础,可支持多领域政府项目的专家组推荐,从而弥补现有方法的空缺。对长期累积的专家评审意见进行知识抽取与知识计量的方法由于得到科学的知识表示而具有合理可行的计算复杂度,这在实验中得到了充分证实。另一方面,我国政府各类项目管理。部门普遍采用专家库的形式实施项目评审,这为本文方法的实施提供了基础环境。最后,本文方法的效果受被评审对象知识概念的完整性与准确性约束,还受到专家评审意见语言规范性等因素的影响,实用中需要对领域知识体系进行不断的补充与完善。后续将进一步研究本文方法在相关领域的应用,不断提升理论研究与实践水平。

参考文献:

[1] 国家发展和改革委员会. 国家电子政务工程建设项目管理暂行办法:发展改革委令第 55 号[Z]. 北京,2007:1-19.
[2] 史宗恺,宗俊峰,王燕,等.《网络评价》理念上的融合与发展

[J]. 清华大学教育研究,2005(S1):23－26.
[3] 沙毓英,张锋,金竟明,等. 中国人性格研究的理论与方法初探[J]. 云南师范大学学报(哲学社会科学版),1993(3):1－16.
[4] DALTON R. Peers under pressur.[J]. Nature,2001,413(6852).
[5] MENGES E. The perils of peer review[J]. Natural areas journal:2021,41(1).
[6] 游庆根,张金隆,马辉民. 基于熵权法的评审专家遴选模型研究[C]//第十五届全国计算机模拟与信息技术学术会议. 长沙:中国优选法统筹法与经济数学研究会计算机模拟分会,2015:8.
[7] 陈媛,樊治平,谢美萍. 科研项目同行评议专家水平的评价研究[J]. 科学学与科学技术管理,2009,30(10):38－42.
[8] 赵千,耿骞,靳健,等. 一种面向主题覆盖度与权威度的评审专家推荐模型研究[J]. 图书情报工作,2017,61(1):80－88.
[9] 王梓森,梁英,刘政君,等. 科研项目同行评议专家学术专长匹配方法[J]. 计算机应用,2021,41(8):2418－2426.
[10] 俞立平,潘云涛,武夷山. 元评价对提高科技评价质量的影响研究[J]. 科学学研究,2012,30(4):512－516.
[11] 朱庆华.《知识元挖掘》评介——兼议情报学的理论研究[J]. 情报科学,2006(12):1899－1902.
[12] 姜春林,张立伟,谷丽,等. 知识单元视角下学术论文评价研究[J]. 情报杂志,2014,33(4):29－34.
[13] 贺颖,邱均平. 基于科学计量的同行评议专家遴选系统模型构建研究[J]. 图书情报工作,2011,55(14):28－31.
[14] 汪建,王裴裴,丁俊. 科技项目专家评审的元评价综合模型研究[J]. 科研管理,2020,41(2):183－192.
[15] 汤建民. 类比源、类比泉和类比知识单元[J]. 科学学研究,2003(5):467－469.
[16] 国家标准委员会，发展改革委，中央网信办. 政务信息系统定义和范围[EB/OL]. [2020－08－20]. https://wenku.baidu.com/view/2088fe392a160b4e767f5acfa1c7aa00b42a9d1d. html?fixfr=FI3ZbuFXkNsHOwhOKiTb8g%253D%253D&fr=income1-wk_go_search-search.
[17] ZW/T 1001－2019. 电子政务系统建设质量要求和测评规范[S]. 北京:国家信息中心，国家电子政务系统质量监督检验中心,2019:1－36
[18] ZW/T 1002－2019. 电子政务系统运行质量要求和测评规范[S]. 北京:国家信息中心，国家电子政务系统质量监督检验中心,2019:1－42.
[19] ZW/T 1003－2019. 政务信息系统整合共享评估规范[S]. 北京:国家信息中心，国家电子政务系统质量监督检验中心,2019:1－16.
[20] ZW/T 1004－2019. 电子政务数据中心技术要求和测评规范[S]. 北京:国家信息中心，国家电子政务系统质量监督检验中心,2019:1－23.
[21] CHOI S,YOON J,KIM K,et al. SAO network analysis of patents for technology trends identification:a case study of polymer electrolyte membrane technology in proton exchange membrane fuel cells[J]. Scientometrics,2011,88(3):863－883.
[22] 华斌,吴诺,李若瑄. 基于知识图谱的电子政务项目评价方法研究与实践[J]. 情报理论与实践,2021,44(2):147－153,146.
[23] 吴建华,魏云. 评价活动的信息原理[J]. 图书情报工作,2008(5):54－57.
[24] 王忠群,皇苏斌,修宇,等. 基于领域专家和商品特征概念树的在线商品评论深刻性度量[J]. 现代图书情报技术,2015(9):17－25.
[25] 孟陶然. 评标专家均衡随机抽取模型[J]. 中国管理信息化,2009,12(15):50－53.

作者贡献说明:

华斌:确定选题,提出论文修改意见;
吴诺:论文修改,框架设计;
贺欣:论文初稿撰写,论文所涉及的计算机实验。

Research on Expert Individuals Multi-Feature Depiction and Expert Group Equilibrium Recommendation

Hua Bin Wu Nuo He Xin
School of Science and Technology, Tianjin University of Finance and Economics, Tianjin 300222

Abstract: [**Purpose/significance**] A recommendation method of e-government project review expert group based on multi-feature depiction of individual experts is proposed. It can improve the consistency level of project evaluation among expert groups. [**Method/process**] Taking the long-term evaluation opinions of individual experts as the data source, knowledge element recognition and emotion polarity acquisition were realized by using opinion mining technology. The domain knowledge structure of experts was constructed and updated dynamically. Statistical analysis was used to describe level of expert knowledge, judging depth, emotional characteristics and domain expertise. This paper described the expert feature based on Scientometrics and recommended a combination of experts. [**Result/conclusion**] The method in this paper focuses on the multi-dimensional feature equilibrium of expert group. It has good pertinence for e-government project evaluation, and achieves good application effects.

Keywords: e-government project management opinion mining knowledge unit measurement expert recommendation